

ELECTRONIC CHECKLISTS: EVALUATION OF TWO LEVELS OF AUTOMATION

Everett Palmer
NASA Ames Research Center
Moffett Field, CA

Asaf Degani
San Jose State Foundation
San Jose, CA

ABSTRACT

Two versions of an electronic checklist have been implemented in the Advanced Concepts Flight Simulator (ACFS) at NASA Ames Research Center. The two designs differ in the degree of pilot involvement in conducting the checklists. One version (manual-sensed), requires the crew to manually acknowledge the completion of each checklist item. The other version (automatic-sensed), automatically indicates completed items without requiring pilot acknowledgement. These two designs and a paper checklist (as a control condition) were evaluated in line-oriented simulation. Twelve air crews from one major air carrier flew a routine, four-leg, short-haul trip. This paper presents and discusses the portion of the experiment that was concerned with measuring the effect of the degree of automation on the crews' performance. It discusses and presents evidence for a potential down side of implementing an electronic checklist that is designed to provide fully redundant monitoring of human procedure execution and monitoring.

INTRODUCTION

Normal checklist procedures are used in the cockpit to provide a final check that the aircraft has been properly configured for each phase of flight. The use of checklists in aviation probably dates from the first accident in which a pilot failed to put the gear down for landing. Unfortunately, problems exist with the use of this apparently simple device.

The improper use of checklists has been cited as a factor in recent aircraft incidents and accidents (NTSB 1988, 1989, 1990). Degani and Wiener (1990) conducted a field study to analyze factors affecting successful and unsuccessful use of normal air carrier checklists. They found that the current paper checklists have a number of design weaknesses. These problems included the lack of a pointer to the current checklist item, the inability to mark skipped items, and difficulties in getting lost while switching

between checklists. The field study on paper checklists identified a number of problems with paper checklists that may be alleviated with the use of an electronic checklist.

Electronic checklists can provide an external memory of pending, completed, and skipped steps. If the electronic checklist can “sense” and display the state of aircraft subsystems, it can provide a redundant check for completion of many cockpit tasks. It may also guard against automaticity in pilot behavior in which the expected value of a display is perceived instead of the actual value. A touch-operated electronic checklist can also use direct-manipulation graphic techniques to aid the pilot in switching from one procedure to another without losing track of partially completed checklists and without getting lost in a bulky paper procedures manual.

Two electronic checklist systems have been implemented in the Advanced Concepts Flight Simulator at NASA Ames Research Center as part of the Aviation Safety/Automation Program. One goal of the program is to develop systems that enhance the ability of the crew to detect and correct errors before they have serious consequences. We refer to these as “error tolerant systems.”

In this paper we will describe the electronic checklist and the line-oriented simulator study that was conducted to evaluate its design. This paper will focus on the influence of machine checking on human checking performance.

CHECKLIST DESIGNS

The electronic checklist systems were designed with the goal of reducing four types of errors that can occur with conventional paper checklists. These are 1) forgetting what the current item is, and thereby inadvertently skipping an item; 2) skipping items due to interruptions and distractions; 3) intentionally skipping an item and then forgetting to return to it; and 4) stating that an item has been accomplished when it was not

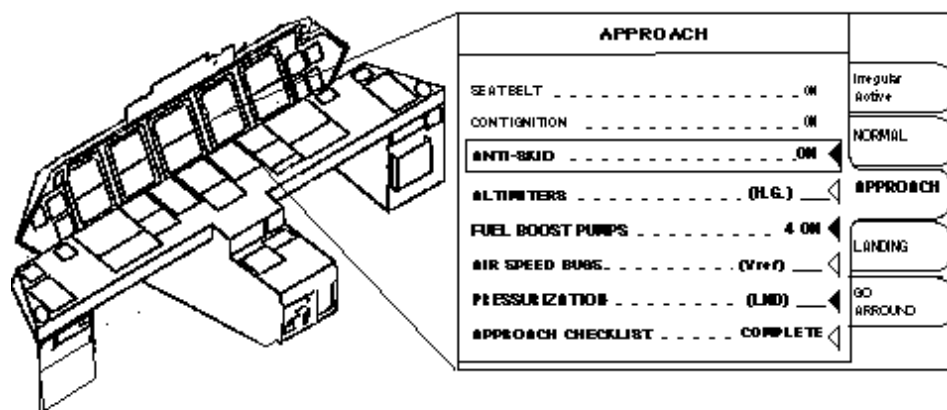


Figure 1. Line drawing of the Advanced Concepts Flight Simulator and a blow up of the manual-sensed checklist. The first two items of the APPROACH checklist are complete.

Line drawings of the simulator cockpit and the “manual-sensed” checklist are shown in Figure 1. The checklist is displayed on the lower portion of the captain’s or first officer’s systems display, while the center display is reserved for engine displays. Each of the

three center electronic displays is touch sensitive. When the checklist display is first called up, all checklist items are colored white and the “current-item box” is situated on the first checklist item. A filled or unfilled triangle, to the right of the checklist item, indicates whether the checklist item is sensed or not. An unfilled triangle indicates that the item cannot be sensed—for example, *Take Off Briefing and Head Count*. A filled triangle indicates an item that can be sensed by the system—for example, *Flap Position and Rudder Trim*. The checklist system automatically displays the relevant synoptic display for the current checklist item in the upper portion of the display.

Manual-Sensed Checklist

Figure 1 shows what the manual-sensed checklist display would look like part way through the execution of the approach checklist. The first two items are colored green and the triangle symbols are removed to indicate their completion. The third item has been manually acknowledged by the pilot, but the checklist software has sensed it to have not been accomplished. The item has been marked as skipped (amber) and the current-item box has not advanced to the next item. The checklist will sense the state of the item only after the pilot touches the display to acknowledge completion. The pilot can either complete the item or skip it by moving the current-item box to the next item.

The last item in every checklist is “*Checklist ... Complete.*” On touching this last item, all skipped or uncompleted checklist items are displayed. The pilot may either return to these items or “override” and continue to the next checklist. The override option was designed to allow the crew to exit a checklist at their discretion even though one or more items were sensed to be incomplete. The checklist design philosophy was to provide reminders, but not to lock out the crew’s control of the situation.

Automatic-Sensed Checklist

In this version, all configuration tasks and actions on the secondary flight displays and overhead panel are still performed by the flight crew, however, the actual operation of the checklist is automated as much as possible. When this checklist is called up, all of the sensed items are checked by the system, and items which are sensed complete are displayed as such (i.e., green and no triangle). Incomplete and unsensed items are displayed as unaccomplished (i.e., white and triangle). These items are accomplished by the pilot in the same manner as in the manual-sensed checklist. If all checklist tasks have been completed and all checklist items are sensed, the only involvement required by the pilot is to manually acknowledge the “*Checklist Complete*” item.

Paper Checklist

The normal paper checklists were printed in standard airline format on a single 8.5 x 11 inch card. The irregular paper checklists were bound in a booklet modeled on Boeing’s Quick Reference Handbook (QRH) format.

The key difference between the two versions of the electronic checklist was whether the electronic checklist system, or the human, checked the “system state” first. We were concerned that pilots who were assigned to the automatic-sensed checklist would be more likely to trust the checklist and not perform their own manual check of system state. The manual-sensed checklist is more time consuming, but it seemed more likely that human

monitoring behavior would be less affected by the presence of machine monitoring. We felt that with the manual-sensed checklist the machine sensing would be more of a true redundant check that the procedure had been done correctly.

EVALUATION METHODOLOGY

The two electronic checklist designs and the control condition (the paper checklist) were evaluated in line-oriented simulation. Checklists are used in all phases of flight and their use is constrained by the actions of many other agents that are external to the cockpit (Degani and Wiener, 1990). We are also convinced that only in line-oriented simulation can one achieve the “suspension of disbelief” that is necessary to induce realistic pilot behavior and crew interactions.

Measuring crew behavior in line-oriented simulation is not an easy assignment. In addition to the usual approach of measuring pilots’ subjective impressions of workload and observer ratings of crew effectiveness, we developed a scenario that included a number of probes of crew monitoring behavior. Our goal was to create a series of small problems (potential pitfalls) in which the crew’s observable behavior would be strong evidence as to their awareness of aircraft configuration. The scenario, included incentives to skip checklist items; several system malfunctions coupled with visual traffic which were timed to interrupt checklist usage; poor checklist sequence; pressures to rush checklist execution; and what we referred to as “configuration probes.” These configuration probes are the focus of this paper.

Configuration Probes

We included three probes to test whether or not the pilots manually checked the state of the system when they performed the checklist task. On leg number two, passing through 9000 feet on the descent to Stockton (KSCA), the scenario software automatically flipped the anti-skid system from “On” to “Off.” On leg number three, descending through 9000 feet on the descent to San Francisco (KSFO), the spoilers were flipped from “Auto” to “Manual.” The checklist called for the anti-skid to be in the “On” position and the spoilers to be in the “Auto” position during these phases of flight. On leg number four, the stabilizer trim was “bumped” to an incorrect setting by the scenario software when the parking brakes were released. The electronic checklists were programmed to wrongly show that these three items had been accomplished when in fact they were not. If the flight crew checked these items while performing the cockpit “flow,” they could easily catch these three discrepancies.

Experiment. Four flight crews were assigned to each of the checklist conditions (paper, manual-sensed, and automatic-sensed). Each flight crew was composed of two pilots, a captain and a first officer. All pilots were current airline pilots and all were rated in a “glass” cockpit aircraft (B-757, 767, 737-300/400, 747-400). Each crew spent two days at the simulator facility. A day and a half were devoted to transition training for the ACFS simulator. The training focused on the basic handling qualities of the airplane and the operation of the aircraft systems in normal and irregular situations. Pilots were instructed and practiced a “flow” for configuring the airplane before calling for and conducting the checklists. Pilots were instructed to run all normal checklists as “check” lists rather than “do” lists. The checklists performed on the ground were done by both pilots in a

challenge - response format. The inflight checklists were called for by the pilot flying and performed silently by the pilot not flying. The mission was flown without the auto-flight system and only conventional VOR navigation was used. On the morning of the second day, the crew completed their training by flying from San Francisco to Sacramento and back.

On the afternoon of day two, the crew flew the experimental mission which consisted of four segments (legs). The flight schedule is shown in Table 1. All normal paperwork for dispatch, weather, maintenance, and flight plan that the crews are accustomed to in their daily line operations were provided. The experiment operator provided all company, ground, dispatch, maintenance, and flight attendant communications and services. An air traffic controller provided ATIS, weather information, clearances, as well as ground, tower, approach, departure, and center communications. ATC communication was provided to other “pseudo” aircraft that were presumably on the same frequency. A “pseudo” pilot, equipped with a voice disguiser, provided all normal radio communications from these “pseudo” aircraft.

FLIGHT #1103	SFO	-	SMK	1700Z	1740Z
FLIGHT #1104	SMF	-	SCK	1845Z	1923Z
FLIGHT #1104	SCK	-	SFO	1950Z	2035Z
FLIGHT #1104	SFO	-	LAX	2120Z	2255Z

Table 1. Flight schedule for the four leg mission.

This experiment was conducted in conjunction with another experiment to evaluate the usefulness of a spatial audio display to aid the visual acquisition of airborne traffic (Begault, 1991). It was felt that the presence of this experiment did not confound the data. Moreover, the out-the-window traffic spotting task added to the realism of the simulation.

Data

The dependent variables were errors, execution time, subjective ratings, TLX workload ratings, inflight observer ratings, and post-flight pilot ratings. This paper only addresses the crew response to the three configuration probes. These probes were designed to test if the pilots were continuing to manually monitor system state when the sensed checklist was also monitoring the systems.

RESULTS

The data in Figure 2 shows that the anti-skid probe was detected by all four of the crews with the paper checklist, by two of the four crews with the manual-sensed checklist, and none of the crews with the automatic-sensed checklist. The spoiler probe was detected by three of the four crews with the paper checklist, one of the crews with the manual-sensed checklist, and none of the crews with the automatic-sensed checklist. All twelve crews detected the stabilizer trim probe either while performing the checklist or during taxi out to the active runway.

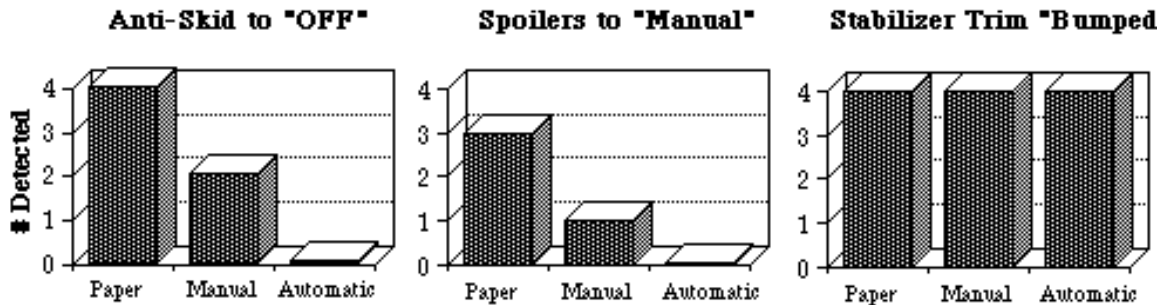


Figure 2. Detection of subsystem failures ("Configuration probes").
Note that there were four crews in each checklist condition.

DISCUSSION

The results illustrate some of the difficulties of adding machine monitoring capability to a system in a desire to increase and enhance its error tolerance. In a fully automated system, a redundant monitoring capability can be added and the increase in overall system reliability can be calculated mathematically. In a human-machine system, human behavior is seldom independent of observable changes in machine functionality. Wiener and Curry (1980) referred to this behavioral phenomena as "Primary/Secondary Task Inversion." In some conditions, operators may begin to rely on the backup system as the primary system.

Take for example the altitude alert installed in some commercial cockpits. This device sounds at about 700 feet before reaching a selected altitude and may become the pilot's primary source of altitude awareness. The crew responses to the three "configuration probes" and our observations of crew behavior during the experiment suggest that both electronic checklist designs encouraged flight crews to not conduct their own checks. Manual checking was largely replaced with machine checking for most inflight checklists. The inflight checklists were run silently by the pilot not flying and the ground checklists were run by both pilots in a challenge - response method. The manual-sensed checklist was designed to encourage manual checking behavior,

but the results show little evidence that it was any more successful in promoting human checking than the automatic-sensed checklist. What appeared to help maintain pilot monitoring was conducting the procedure in the challenge - response method.

CONCLUSION

When an automatic monitoring capability is added to the system, it is a mistake to assume that human behavior will be unaffected. It is likely that machine monitoring of the system will be largely replace, not add to, the human monitoring of system state. The loss of human monitoring is not necessarily a drawback, as the combined human and machine monitoring may be more reliable than the unaided human. What designers can not assume is that a machine monitoring system provides true redundancy to human monitoring. An additional concern is the potential loss in human situation and system awareness that may be associated with a decrease in human monitoring.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge David Pepitone's assistance in managing this experiment, Michael Smiley's help in designing the scenario, Anna Okseniuk's help in programming the checklists, and the rest of the programming and operational staff of the Man-Vehicle Systems Research Facility for their assistance in conducting this simulation. Asaf Degani was supported by NASA Ames Research Center Grant NCC2-327 to the San Jose State University.

REFERENCES

- Begault, D. R., Stein, N. E., and Loesche, V. (1991). *Advanced Audio Applications in the NASA-Ames Advanced Cab Flight Simulator*. Submitted for publication as a NASA Technical Memorandum.
- Degani, A., & Wiener, E. L. (1990). *The human factors of flight-deck checklists: The normal checklist* (NASA contractor report 177549). Moffett Field, CA: NASA Ames Research Center.
- National Transportation Safety Board. (1988). *Northwest Airlines. DC-9-82 N312RC, Detroit Metropolitan Wayne County Airport. Romulus, Michigan. August 16, 1987* (Aircraft accident report, NTSB/AAR-88/05). Washington, DC.
- National Transportation Safety Board. (1989). *Delta Air Lines, Boeing 727-232, N473DA. Dallas-Fort Worth International Airport, Texas. August 31, 1988* (Aircraft accident report, NTSB/AAR-89/04). Washington, DC.
- National Transportation Safety Board. (1990). *USAir, Inc., Boeing 737-400, N416US. LaGuardia Airport. Flushing, New York. September 20, 1989* (AAR-90/03). Washington, DC.
- Wiener, E. L., & Curry, R. E. (1980). Flight-deck automation: Promises and problems. *Ergonomics*, (23), 995-1011.